Relativity in Modern Astrometry and Celestial Mechanics — An Overview

S. Klioner

Lohrmann Observatory, Dresden Technical University, 01062 Germany

Abstract. After a short introduction into experimental foundations of general relativity, an overview of modern relativistic modelling of astronomical observations is given. The principal relativistic effects in various kind of observations are reviewed, and the problems for increasing accuracy of the astrometric catalogs due to general relativity are formulated.

1. Introduction

The rapid increase of observational accuracy of astronomical observations makes it indispensable to use general relativity to model observational data. Recognizing this fact, the IAU adopted in 1991, and several times afterwards refined, a relativistic framework to be used for the model. This paper gives an overview of the general scheme of the relativistic modelling of astronomical data and the principal observable relativistic effects.

2. Experimental foundations of general relativity

Although Einstein's general relativity is by no means the only theory of gravitation, it seems to be the most simple one among the theories successfully passing all the observational tests. Let us first briefly review the experimental foundations of general relativity. A very detailed overview of the modern experimental foundations of gravitational physics can be found in Will (1998). The basic principle of the theory is called the Einstein Equivalence Principle. This principle consists of three parts. The first is the Weak Equivalence Principle stating that the masses on the both sides of the Newtonian gravitational law,

$$m_{iner}\,\ddot{r}^i = -G\,m_{grav}\,M\,r^i\,/\,r^3,$$

exactly coincide $m_{iner} \equiv m_{grav}$. This has been tested in many different experiments with a precision of $|\delta m|/m < 5 \cdot 10^{-13}$. The second part is the local Lorentz invariance, that is, the postulate of special relativity stating that the light velocity in vacuum, c, is constant in any inertial reference system. This has been tested at a level of $|\delta c|/c < 10^{-21}$. The third part is the local positional invariance which can be tested by measuring the gravitational red shift

$$\Delta\nu\,/\,\nu = (1 + \alpha)\,c^{-2}\,\Delta U,$$

where $\alpha = 0$ in general relativity. Again, a number of different experiments have proved that $|\alpha| < 2 \cdot 10^{-4}$.

However, the Einstein Equivalence Principle does not necessarily imply general relativity. There exists a number of alternative metric theories of gravity compatible with that Principle, but producing different observable effects. Therefore, the observable predictions of general relativity should be further tested. The most important tests of general relativity are: (1) the Mercury perihelion shift measured mainly by Mercury radar ranging with a precision of 0.0012, (2) the gravitational light deflection measured by Hipparcos with a precision of 0.003, (3) the gravitational signal retardation (Shapiro effect) measured by Viking radar ranging with a precision of 0.002, (4) the differential Shapiro effect measured by geodetic VLBI observations with a precision of 0.0003, (5) the geodetic precession observed by Lunar Laser Ranging (LLR) with a precision of 0.007, and finally (6) the Nordtvedt effect also measured by LLR with a precision of 0.001. This allows one to conclude that the main observable predictions of general relativity have been tested with a precision of $\sim 0.1\%$. A special theoretical framework called Parametrized Post-Newtonian (PPN) formalism can be used here to describe the experimental results as constraints on the numerical values of the PPN parameters β, γ, *etc.* that can often be found in the literature.

Future space missions will give us the opportunity to test general relativity with even better accuracy. It is sufficient to mention here the STEP project intended to test the Weak Equivalence Principle with a precision of 10^{-18}, FAME which will be able to test the gravitational light deflection with a precision of 10^{-4} or slightly better, GAIA and SIM which will measure the same effect with a precision of 10^{-6}, and Gravity Probe B designed to measure the geodetic precession with a precision of $5 \cdot 10^{-5}$.

3. General scheme of relativistic modelling of astronomical observations

The general scheme of relativistic modelling of astronomical observational data is represented in Fig. 1. Starting from the general theory of relativity one should define at least one relativistic 4-dimensional reference system covering the region of space-time where all the processes constituting particular kinds of astronomical observations are located. Typical astronomical observations depicted in Fig. 2 consists of four constituents to be modelled. The equations of motion of both the observed object and the observer relative to the chosen reference system should be derived and a method to solve these equations should be found. Typically the equations of motion are second-order differential equations, and numerical integration can be used to solve them. What the observer observes is the electromagnetic signals come from the object. The equations of motion of the object and the observer and the equations of light propagation enable one to compute positions and velocity of the object, observer and the photon (light ray) with respect to the particular reference system at a given moment of coordinate time, provided that the positions and velocities at some initial epoch are known. However, these positions and velocities obviously depend on the reference system used. On the other hand, the results of observations cannot depend on the reference system used to theoretically model the observations. Therefore, one more step of modelling is needed: the relativistic description of
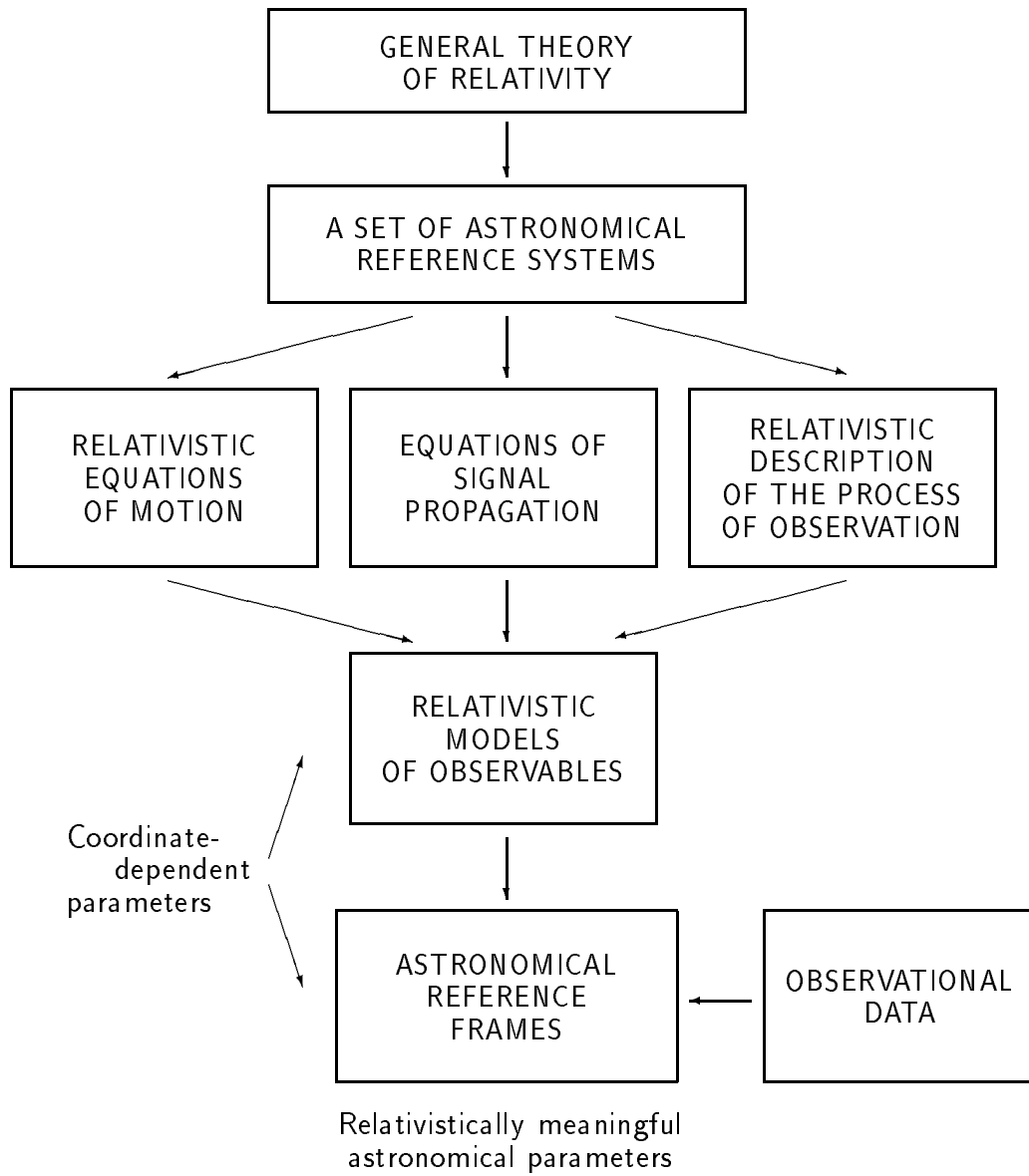
Figure 1.   General principles of relativistic modelling of astronomical observations.
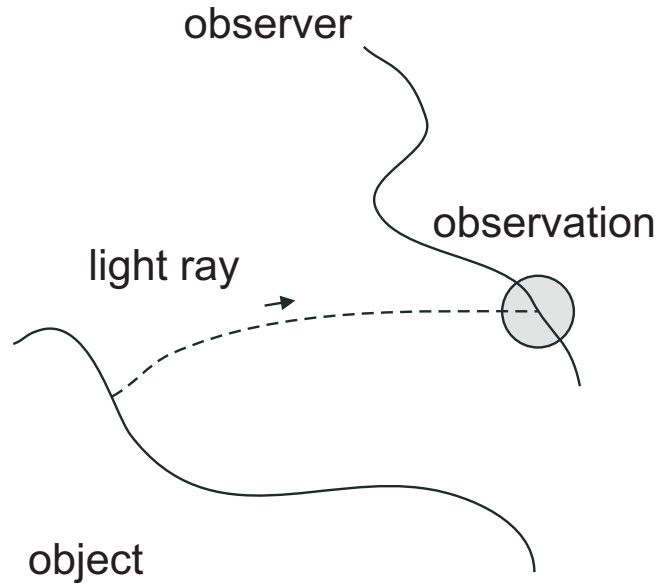
Figure 2.   Four parts of an astronomical event: 1) motion of the observed object, 2) motion of the observer, 3) propagation of an electromagnetic signal from the observed object to the observer, 4) the process of observation.

the process of observation. This allows one to compute a coordinate-independent theoretical prediction of the observables starting from the coordinate-dependent position and velocity of the observer and, in some cases, the velocity of the electromagnetic signal at the point of observation.

The mathematical techniques used to derive the equations of motion of the observed object and the observer, and the equations of light propagation and those describing the process of observation in the framework of general relativity are well known. These three parts can now be combined into relativistic models of observables. The models give an expression for each observable under consideration as a function of a set of parameters. These parameters should be fit to observational data to produce astronomical reference frames, which represent a set of best estimates of certain parameters appearing in the relativistic models of observables. For example, the International Celestial Reference Frame (ICRF) represents a catalog of coordinates of extragalactic radio sources with respect to the Barycentric Celestial Reference System (BCRS), which is a well-defined relativistic 4-dimensional reference system.

It is very important to understand at this point that the relativistic models contain parameters which are defined only in the chosen reference system and are thus coordinate-dependent. A good example of such coordinate-dependent parameters are the coordinates and velocities of various objects at some initial epoch. On the other hand, from a physical point of view any reference system covering a region of space-time under consideration can be used to describe physical phenomena within that region, and we are free to choose the reference system to be used to model the observations. However, some reference systems, in which mathematical description of physical laws is simpler than in others, are

more convenient for practical calculations. Therefore, one can use the freedom to chose the reference system to make the parametrization as convenient and reasonable as possible (*e.g.*, one prefers the parameters to have simpler time-dependence).

These convenient reference systems are called local reference systems. The principal physical idea of local reference systems are illustrated in Fig 3. Currently there exist two well-developed theoretical formalisms to solve the problem of local reference systems in General Relativity: the Brumberg-Kopeikin formalism (see, *e.g.*, Brumberg, 1991; Kopeikin, 1991; Klioner, Voinov, 1993) and the Damour-Soffel-Xu formalism (Damour *et al.*, 1991, 1992, 1993). The results of these formalisms prove that a local reference system of a massive extended body in general relativity has two properties: (A) the gravitational field of external bodies is represented in the form of tidal potentials being $\mathcal{O}(\mathbf{X}^2)$, where X^i are local coordinates; (B) the internal gravitational field of the body coincides with the gravitational field of a corresponding isolated source provided that the tidal influence of the external matter is neglected. The IAU recommendations (both IAU (1991), and the new proposal) are based, in fact, on the results of these two formalisms. Various parameters of the models defined in such local reference systems usually can be reasonably interpreted at some level of accuracy neglecting the fact that the parameters are defined in one particular reference system and have, generally speaking, no physical meaning. For example, at a sub-millimeter level of accuracy the local geocentric coordinates of Earth-bound stations derived from VLBI, GPS, SLR and other techniques can be interpreted directly in the framework of Newtonian geodynamics.

For a system of N massive bodies $N + 1$ reference systems are considered: one global reference system covering the whole space-time (t, x^i) and one local reference system (T, X^a) for each of the bodies. Each reference system is defined by the structure of its metric tensor. Thus, for the global reference system (t, x^i) the metric tensor reads.

$$
\begin{aligned}
g_{00} &= -1 + 2\,c^{-2}\,w(t, \mathbf{x}) - 2\,c^{-4}\,w^2(t, \mathbf{x}) + \mathcal{O}(c^{-5}), \\
g_{0i} &= -4\,c^{-3}\,w^i(t, \mathbf{x}) + \mathcal{O}(c^{-5}), \\
g_{ij} &= \delta_{ij}\left(1 + 2\,c^{-2}\,w(t, \mathbf{x})\right) + \mathcal{O}(c^{-4}).
\end{aligned}
$$

Here w and w^i are the post-Newtonian gravitational potentials. The metric tensors of each local reference system have the same functional form, but with local post-Newtonian potentials W and W^a. The theory also gives the 4-dimensional coordinate transformations between any local reference system (T, X^a) and the global one (t, x^i). The explicit form of the metric tensors and the coordinate transformations are given in the papers devoted to the Brumberg-Kopeikin and the Damour-Soffel-Xu formalisms cited above.

4. Principal relativistic effects

Let us now review the most important relativistic effects appearing in typical astronomical observations. The effects can be divided into the effects in relativistic equations of motion (translational and rotational), the effects in light propagation and the effects in the conversion to observables.
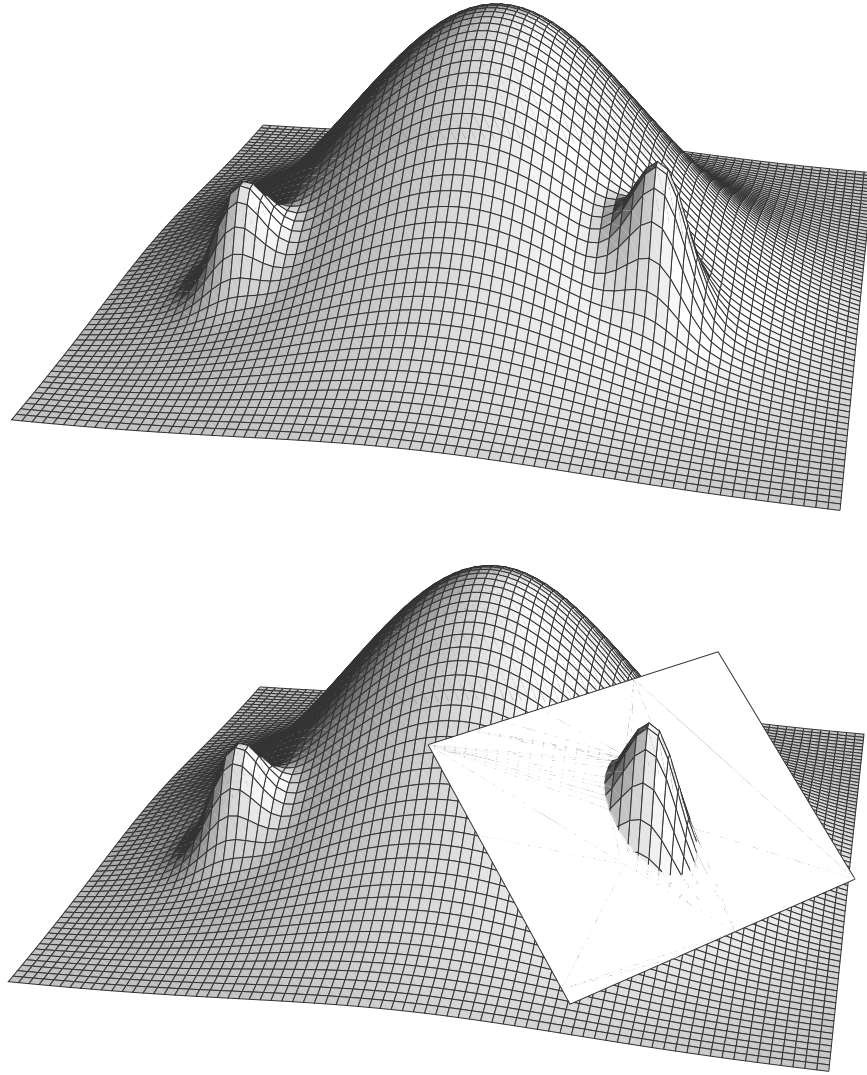
Figure 3.   The basic idea of a local reference system can be illustrated
with the help of these two pictures. The upper picture is a graphical
representation of the curved space-time produced by a system of one
massive central body (say, the Sun) and two less massive ones (say, two
planets). If the system as a whole is to be described, all three sources
of the gravitational field should be described uniformly. However, if
only a relatively small vicinity of one of the bodies is to be considered,
one can introduce a more convenient reference system called a local
reference system. Since the background curvature due to the two other
bodies changes very slowly within that region, a local reference system
(see, lower picture) can be introduced, where only the curvature due
to the body considered is effectively seen. It is intuitively clear that
the background curvature of the space-time still appears in the local
reference system, but in tidal form.

4.1. Translational motion

The principal relativistic effects in the translational motion are contained in the so-called Einstein-Infeld-Hoffmann (EIH) equations of motion of N gravitating bodies, whose gravitational fields can be described by their masses M_A only:

$$\ddot{\mathbf{x}}_A = - \sum_{B \neq A} G\, M_B \frac{\mathbf{x}_A - \mathbf{x}_B}{|\mathbf{x}_A - \mathbf{x}_B|^3} + \frac{1}{c^2} \mathbf{F}_{pN}(M_B, \mathbf{x}_B, \dot{\mathbf{x}}_B) + \mathcal{O}(c^{-4}).$$

Deriving these equations requires all the terms in the metric tensor specified above. Various parts of these equations represent: (1) perihelion shifts (~ 43 $''$/cty for Mercury, ~ 10 $''$/cty for Icarus, *etc.*); (2) geodetic precession (~ 2 $''$/cty for the lunar orbit); (3) various periodic relativistic effects (important mostly for LLR and binary pulsar timing observations). Further effects not contained in the EIH equations are the effects due to rotation of the bodies (Lense-Thirring or gravitomagnetic effects) and those due to non-sphericity of the gravitating bodies. These additional effects are marginal for the current accuracy of LLR and SLR. Depending on the problem, both a global barycentric reference system (*e.g.*, for the motion of planets) and a local geocentric reference system (*e.g.*, for the motion of an Earth satellite) can be used.

4.2. Rotational motion

Adequate relativistic description of rotational motion (*e.g.*, of the Earth) requires the use of the local reference system of the body under study. The post-Newtonian equations of motion read

$$d^2\, S^i\, /\, dt^2 = L^i_N + c^{-2}\, L^i_{pN} + \mathcal{O}(c^{-4}),$$

where S^i is the relativistic spin of the body, L^i_N and L^i_{pN} are the Newtonian and post-Newtonian torques, respectively (see, Klioner and Soffel, 1998 and reference therein for further details). Again the full post-Newtonian metric tensor is required to derive these equations. The most important effects in the Earth's rotation are the geodetic precession ($\sim 1.914\,''$/cty) and nutation (~ 153 μas), and the effects due to L^i_{pN} ($\sim 1\mu$as for the Earth and up to 1 mas for the Moon).

4.3. Light propagation

The post-Newtonian equations of light propagation read

$$\mathbf{x}(t) = \mathbf{x}_0(t) + c\, \boldsymbol{\sigma}\, (t - t_0) + c^{-2}\, \mathbf{S}_{pN}(t) + c^{-3}\, \mathbf{S}_{1.5pN}(t) + c^{-4}\, \mathbf{S}_{ppN}(t),$$

where $\mathbf{S}_{pN}$ and $\mathbf{S}_{ppN}$ are the post-Newtonian and post-post-Newtonian effects, and $\mathbf{S}_{1.5pN}$ is the additional effect induced by the motion of gravitating matter (*i.e.*, by translational and rotational motion of the gravitating bodies). The terms of order of c^{-2} in both g_{00} and g_{ij} are required to derive $\mathbf{S}_{pN}(t)$. The terms of order of c^{-4} in both g_{00} and g_{ij} are needed for $\mathbf{S}_{ppN}$, and the terms c^{-3} in g_{0i} for $\mathbf{S}_{1.5pN}$. The principal observable effects in light propagation are (1) the gravitational light deflection (amounting to $1\overset{''}{.}75$ for a light ray grazing

the Sun) and (2) the gravitational signal retardation (the Shapiro effect). This effect amounts to ∼ 240 μs for the radar ranging of Venus in upper conjunction.

For the planned astrometric space missions it is obviously important to model the gravitational light deflection with high accuracy. Table 1 illustrates the post-Newtonian and post-post-Newtonian gravitational light deflection due to the Sun, the Moon and the 8 major planets. Beside the effects shown in Table 1, the light deflection $\mathbf{S}_{1.5pN}$ due to translational and rotational motion of the gravitating bodies may amount to ∼ 1μas. A detailed discussion of various effects can be found in (Klioner and Kopeikin, 1992). It is interesting to note that a number of smaller bodies should also be taken into account. For a spherical homogeneous body with density ρ, the light deflection is larger than δ if its radius

$$ R \geq \left(\frac{\rho}{1\,\mathrm{g/cm^3}} \right) \cdot \left(\frac{\delta}{1\,\mu\mathrm{as}} \right) \cdot 650\,\mathrm{km}. $$

Therefore, at a level of 10 μas one should account for Io (30 μas), Europa (19 μas), Ganymede (35 μas), Callisto (28μas), Titan (32 μas) and Triton (20 μas). The influence of the Galilean satellites attains 1 μas at the angular distance of 11–32″. At a level of 1 μas one should additionally account for Pluto (7 μas), Charon (4μas), Iapetus, Rea, Dione, Ariel, Umbriel, Titania, Oberon and Ceres (1–3 μas).

body	pN		Q		ppN	
Sun	1.75″	(180°)	∼1		11	(53′)
Mercury	83	(9′)	—		—	
Venus	493	(4.5°)	—		—	
Earth	574	(178°/124°)	—		—	
Moon	26	(9°/5°)	—		—	
Mars	116	(25′)	—		—	
Jupiter	16300	(90°)	240	(3′)	—	
Saturn	5800	(18°)	95	(51″)	—	
Uranus	2100	(72′)	25	(6″)	—	
Neptune	2600	(51′)	10	(3″)	—	

Table 1.   Various components of the gravitational light deflection: pN and ppN are the post-Newtonian and post-post-Newtonian effects due to the spherically symmetric field of each body, Q are the effects due to non-sphericity of the bodies. The symbol "—" means that the effect is smaller than 1 μas. The values in parentheses are maximal angular distances between the body and the source at which the corresponding effect still attains 1 μas. The observer is supposed to be in the vicinity of the Earth. For the Earth and Moon two estimations are given: for a geostationary satellite and for a satellite at a distance of 10^6 km from the Earth.

4.4. Conversion to observables: proper direction

As mentioned above, the conversion of the coordinate-dependent quantities into coordinate-independent observables is an important part of the relativistic models. Mathematically the coordinate-independent quantities are scalars, and special mathematical techniques are known to perform the conversion in each particular case. One of the most important application of this conversion procedure is a conversion of the coordinate angle between two sources φ_{coord} into the corresponding observable angle φ_{obs}:

$$\varphi_{\mathrm{obs}} = \varphi_{\mathrm{coord}} + c^{-1}\,S_1 + c^{-2}\,S_2 + c^{-3}\,S_3 + \mathcal{O}(c^{-4}),$$

where $S_i(\varphi_{\mathrm{coord}}, \mathbf{v_o})$ are the aberrational effects of the i-th order, and $\mathbf{v_o}$ is the coordinate velocity of the observer. S_1 is the classical Newtonian aberration. The explicit expression for S_i can be found, *e.g.*, in (Klioner, Kopeikin, 1992). For $|\mathbf{v_o}| \sim 40$ km/s one has $|c^{-1}S_1| \sim 28''$, $|c^{-1}S_2| \sim 3.6$ mas and $|c^{-1}S_3| \sim 1\,\mu$as. Note also that the higher-order aberrational effects S_i, $i > 1$ are nonlinear with respect to the velocity $\mathbf{v_o}$ and cannot be divided into pieces like "annual" and "diurnal" aberrations. Let us also note a challenge for the planned space missions: in order to calculate $c^{-1}\,S_1$ with an accuracy of 1 μas, the velocity of the satellite $\mathbf{v_o}$ must be known with an accuracy of 10^{-3} m/s.

4.5. Conversion to observables: proper time

Another important case is the conversion of intervals of the coordinate time t into the corresponding intervals of the proper time τ of an observer. The general form of the conversion reads

$$d\tau\,/\,dt = 1 + c^{-2}\,A_{pN} + c^{-4}\,A_{ppN} + \mathcal{O}(c^{-5}),$$

where A_{pN} and A_{ppN} are the post-Newtonian and post-post-Newtonian terms, respectively. The explicit form of these two functions depends on the metric tensor. In order to compute for A_{pN}, the terms of order of c^{-2} in g_{00} are needed, while the terms of order of c^{-4} in g_{00}, c^{-3} in g_{0i}, and c^{-2} in g_{ij} are required to compute A_{ppN}. Typically in the solar system $|c^{-2}\,A_{pN}| \sim 10^{-8}$ and $|c^{-4}\,A_{ppN}| \sim 10^{-16}$. Therefore, the post-post-Newtonian terms A_{ppN} are needed to cope with modern high-precision clocks, the stability of which may attain 10^{-18} in the near future. Let us also mention that in order to compute the drift rate of a clock on the Earth with a precision of 10^{-18}, its height above the geoid should be known with an accuracy of ~ 9 mm and the tidal deformations of the Earth's surface should be known with even better accuracy. This probably means that the future TAI-like time scales having a precision of 10^{-18} or better will be based on space-based clocks.

5. Increasing astrometric accuracy and relativity

General relativity seems to bring several complications for increasing astrometric accuracy below some level. The first one to mention is the so called weak microlensing which is simply the gravitational deflection of the light coming from a distance source produced by the gravitational field of a visible or invisible object

situated between the observed source and the observer. For an object of mass M at a distance r from the solar system and at an angular distance φ from the source assumed to be much further than the deflecting object the deflection angle $\delta\varphi$ reads

$$\delta\varphi \approx (M \,/\, M_{\mathrm{sun}}) \cdot (r \,/\, 1\mathrm{pc})^{-1} \cdot (\tan\varphi \,/\, \tan 1'')^{-1} \cdot 8000 \,\mu\mathrm{as}.$$

For example, for $M = 0.1\,M_{\mathrm{sun}}$, $r = 20\mathrm{pc}$ and $\varphi = 5''$, the effect of lensing is $\delta\varphi = 8\,\mu\mathrm{as}$. The problem is not the gravitational deflection itself, but its time dependence (otherwise the constant changes in the positions of the sources could be absorbed in the "observed" positions as in the case of secular aberration). In the numerical example above, the change of the angular distance by $\Delta\varphi = 1''/\mathrm{yr}$ produces a change in the gravitational deflection angle $\Delta\delta\varphi = 1.3\,\mu\mathrm{as}/\mathrm{yr}$. It is clear that weak microlensing leads to unpredictable (and in some cases undetectable) errors in proper motions of the sources. A statistical analysis of this effect is given in Sazhin *et al.* (1998) and references therein.

Further effects to be mentioned are the gravitational lensing on stochastic (unpredictable) gravitational waves (Pyne *et al.*, 1996; Kopeikin *et al.*, 1999), and the necessity to consider cosmological models to interpret the distance data (1μas accuracy of parallaxes allows one to measure the distance to the objects as far as 1 Mpc away from the solar system. See, Kristian and Sachs (1965) for a discussion of the astrometric consequences of cosmology).

References

Brumberg, V.A., 1991, *Essential Relativistic Celestial Mechanics*, Adam Hilger, Bristol.

Damour, T., Soffel, M., Xu Ch., 1991, *Phys. Rev. D*, **43**, 3273; 1992, **45**, 1017; 1993, **47**, 3124.

IAU 1991, *Transactions of the International Astronomical Union*, Vol. XXIB, Kluwer, Dordrecht.

Klioner, S.A., Kopeikin, S.M., 1992, *Astron. J.*, **104**, 897.

Klioner, S.A., Voinov, A.V., 1993, *Phys. Rev. D*, **48**, 1451.

Klioner, S.A., Soffel, M., 1998, *Highlights of Astronomy*, 11A, 173.

Kopeikin, S.M., 1991, in "Itogi Nauki i Tekhniki", edited by M. Sazhin, Nauka, Moscow, 87 (in Russian).

Kopeikin, S.M., Schäffer, G., Gwinn, C.R., Eubanks, T.M., 1999, *Phys. Rev. D*, 59, 084023.

Kristian, J., Sachs, R.K., 1965, *Astrophys. J.*, **143**, 379.

Pyne, T., Gwinn, C.R, Birkinshaw, M., Eubanks, T.M., Matsakis, D.N., 1996, *Astrophys. J.*, **465**, 566.

Sazhin, M.V., Zharov, V.E., Volynkin, A.V., Kalinina, T.A., 1998, *Mon. Not. RAS*, **300**, 287.

Will, C.M., 1998, in Lecture notes from the 1998 Slac Summer Institute of Particle Physics (available at `http://xxx.lanl.gov/ps/gr-qc/9811036`).